DOCUMENT RESUME

ED 358 731 FL 021 298

AUTHOR Kanno, Yasuko
TITLE Do Formulaic Utterances Cease to Be "Chunks" When They Are Analyzed?
PUB DATE 93
NOTE 20p.; For the complete volume, see FL 021 293.
PUB TYPE Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Information Analyses (070) -- Journal Articles (080)
JOURNAL CIT MITA Working Papers in Psycholinguistics; v3 p75-92 1993

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Context Effect; Foreign Countries; *Language Processing; Language Research; *Linguistic Theory; *Oral Language; *Second Language Learning; *Syntax

ABSTRACT

Many second-language learners in their early stages of development are known to make an extensive use of prefabricated formulae. These formulae are extracted holistically from the input and memorized by rote. Learners can learn to use expressions that are far beyond their current knowledge of syntax and vocabulary, by guessing their meaning from the contextual cues. The formulae that learners use therefore usually sound more fluent and linguistically advanced than their creative speech (i.e. utterances they generate using syntactic rules). A major dispute over the role of formulaic utterances in the 1970s was whether they led to creative language or they were a dead-end. This paper argues that formulaic utterances remain as units or "chunks" in the learner's lexicon, even after their syntactic structure becomes apparent to the learner, as long as they serve some purpose in economizing processing energy in sentence production. That is, the lack of syntactic analysis is not a defining character of such formulae. Data from various research is used to support the argument and demonstrate how it can explain some second language acquisition phenomena better than extant theories that have been widely used to explain them. Theoretical implications are then explored. (VWL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED358731

Do Formulaic Utterances Cease to Be "Chunks" When They Are Analyzed?

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Yukio Otsu

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

EJ 521 998

MITAWPP 3 (1993) 75-92

Do formulaic utterances cease to be "chunks" when they are analyzed?

Yasuko Kanno

Many second language learners in their early stages of development are known to make an extensive use of prefabricated formulae. These formulae are extracted holistically from the input and memorized by rote. Learners can learn to use such expressions as What's that?, Don't do that, and I'm finished, which are far beyond their current knowledge of syntax and vocabulary, by guessing their meaning from the contextual cues. The formulae that learners use therefore usually sound far more fluent and linguistically advanced than their creative speech (i.e. utterances that they generate using syntactic rules).

Researchers have called such formulas by various names: 'formulaic utterances' (Wong Fillmore 1976), 'prefabricated routines and patterns' (Hakuta 1974, Krashen and Scarcella 1978), 'speech formulas' (Peters 1977, 1983), 'prefabricated chunks' (Widdowson 1989) and 'lexicalized sentence stems' (Pawley and Syder 1983). However, they all refer to the same thing. What is more significant here is the definition of such formulaic utterances. Essentially, the definition, which is implicit in various researchers' work, is two-fold: 1) that formulae are extracted from the input and stored in the lexicon *as units*; and 2) that the learner does not know their internal structure.

A major dispute over the role of formulaic utterances in 1970's was whether they lead to creative language or they are a dead-end. More precisely, Wong Fillmore (1976, 1979), Clark (1974) and Peters (1983), on the one hand, argued that formulae that the learner has acquired constitute the data source from which syntactic rules are developed. On the other hand, Krashen and Scarcella (1978) maintained that formulae and rule-formation are developed separately. The controversy has not been settled and in fact it seems to have been dropped over the past decade. Although this topic is theoretically quite interesting and is relevant to my argument, it is not my main concern here. My central argument concerns one issue that both parties seem to agree on: that when the learner becomes aware of the internal structure of formulaic utterances, whether by comparing a few similar

formulae or by applying independently developed syntactic rules, they lose their status as unitary items in the learner's lexicon. This argument is basically based on the principle of parsimony rather than a psychologically plausible model. And this is a point where I depart from both parties.

I will argue that formulaic utterances remain as units, or "chunks", in the learner's lexicon, even after their syntactic structure becomes apparent to him/her, as long as they serve some purpose in economizing processing energy in sentence production. That is, lack of syntactic analysis is not a defining character of such formulae. I will draw on various researchers' data on SLA, including my own[1], which support my argument and demonstrate how it can explain some SLA phenomena better than the extant theories that have been widely used to explain them. I will then go on to explore some theoretical implications.

A starting point of creative language or a dead-end?

There has been much dispute over the role of formulaic expressions in SLA. Clark (1974), Wong Fillmore (1976, 1979) and Peters (1977, 1983) argue that these expressions do get analyzed, first partially and then fully, and will eventually lead to syntactic rules. Peters (1983) assertion is representative of this position: "socially relevant formulaic speech [is] *not* a dead end, but, [leads], through a documentable process of formulaic breakdown, first to formulaic frames with slots and eventually towards analysis into the conventional lexical items and syntactic patterns of the language"(p. 13, original emphasis). On the other hand, Krashen and Scarcella (1978) are of a very different opinion. They claim that the formulae and the syntactic rule formation are developed in different parts of the brain and that therefore there is no interface between them: "prefabricated routines may evolve into patterns[2], but at the same time, independently, the creative construction process develops"(p. 284).

[1] My data come from the participant observation that I conducted from October 1991 to February 1992 in an ESL class of a public elementary school in Toronto. Six children I observed ranged from 7 to 11 in age, and came from various countries: Bulgaria, China, Korea, Sri Lanka, Iran and Rumania. Some of them had immigrated to Canada in September 1991 and others a few months later, but all came to this class immediately after their arrival. These children take regular classes with Canadian students but have one hour of ESL everyday. I went to the ESL class once every one or two weeks and recorded their spontaneous utterances among other things. All the names of the children are pseudonyms.

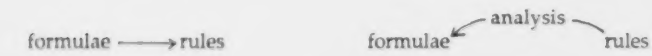
[2] Krashen and Scarcella (1978) distinguish'prefabricated routines' and 'prefabricated patterns'. The former are memorized wholes while the latter are partly fixed expressions with open slots in which appropriate words or expressions can be inserted.

It is important to clarify exactly where the two positions differ, since much confusions seems to have arisen from simple misinterpretation of each claim. The only difference between the two positions concerns the question whether formulaic utterances become the data source from which syntactic rules are derived or not. Wong Fillmore, Clark, and Peters maintain that they do while Krashen and Scarcella argue that they do not. A quotation from Wong Fillmore's doctoral dissertation (1976) is illustrative of her position concerning this point:

> The analyses the learner perform are on those things which are most available to them--the well-practiced and familiar expressions they find in their own speech repertories. How much more reasonable this seems than to assume that the language learner can somehow apprehend the fast-fading message produced by someone else, figure out what it means and how it is put together, and then relate it to similar utterances he has heard (quoted in Peters 1983, p. 14).

In contrast, as it is clear from the quotation above, Krashen and Scarcella believe that syntactic rules are developed elsewhere. Note that they never said that prefabricated routines and patterns are permanently immune to analysis, as some of their opponents seem to interpret their position. In fact, they make their position on this point clear: "in some situations propositional language may 'catch up' with automatic speech; that is, the language acquisition process may 'reanalyze' patterns and routines as creative construction" (1978, p. 284). In other words, they believe that syntactic rules are developed without any reference to the formualic expressions but once they are acquired, they may be applied to analysed the formulae (Figure 1):

Figure 1. The development of syntactic rules.



There is empirical data to support each of the two positions. First of all, Wong Fillmore (1976) has an abundant supply of data that document how hypotheses

about syntactic rules are developed by comparing similar formulae. For example, Nora, the most advanced learner in her study, had the following two prefabricated sentences, among others, at her disposal:

(1) I wann play 'wi dese.
(2) I don' wanna do dese. (quoted in Krashen and Scarcella 1978, p.293)

By comparing (1) and (2), she realised that the constituents after wanna can be exchanged, thus yielding:

(3) I wanna do dese.
(4) I don' wanna play dese. (Ibid.)

In other words, having compared these two formulaic utterances, Nora learned that I wanna and I don' wanna can be followed by a VP. This data clearly show that a syntactic rule was derived from formulaic utterances.

On the other hand, Krashen and Scarcella (1978) take Brown and Hanlon's first language acquisition data (1970) as an example of the rule formation that is developed independently and is later applied to formulaic expressions for analysis. Their subjects, Adam, Eve, and Sarah, acquired such wh-questions as What's that? and What doing? as formulae, an influence from their parents, who used certain wh-questions frequently. What happened after this is the crucial point:

> When, much later, the children began to produce all manner of *wh* questions in the preposed form (such as *What he wants*), it was interesting to note that *What's that?* and *What are you doing?* were not at first reconstrued in terms of the new analysis. If the children had generated the sentences in terms of their new rules, they ought to have said *What that is?* and *What you are doing?* but instead, they, at first, persisted with the old forms . . . . We suggest that any form that is produced with very high frequency by parents will be somehow represented in the child's performance even if its structure is far beyond him . . . . Extensive use of such an unanalyzed or mistakenly analyzed fragments probably protects it, for a time, from reanalysis when the structure relevant to it is finally learned. (p. 51)

If the rule of pre-posing had been formulated on the basis of such formulae as

What's that? and What doing?, then these two expressions would have been the *first* to be affected by the rule. However, that was not the case. On the contrary, these formulae remained unaffected even when other wh-questions were preposed. This, then, is a clear example of a syntactic rule formulated independently.

In the light of both positions having convincing evidence, how can we resolve this debate? My proposal is a rather unexciting one: they are both right. I fail to see why some rules cannot derive from formulaic utterances while others from somewhere else. Language learners, whether it is L1 or L2, use multiple strategies simultaneously, including ways of formulating hypotheses about syntactic rules and testing them. Then, does it not make more sense to suppose that language learners derive rules from multiple sources? In short, these two positions on the role of formulae in SLA are not mutually exclusive; each of them highlights a different strategy for rule-formation.

Syntactic analysis and the status of formulae

Hakuta (1974), at the end of his article, poses an important question for further research: "do prefabricated patterns whose internal structure is finally perceived remain as convenient short-cut routes to production or are they simply discarded, never to be employed again?" (p. 296). Although their positions on the role of formulaic utterances in the formation of syntactic rules are different, various researchers all seem to agree on this point: that once the memorized expressions are analyzed, they lose their status as single units in the lexicon. Krashen and Scarcella (1978) note that "'[formulaic utterances] may be a *temporary* strategy for the performer to outperform his analytic competence, to solve certain communication problems that his creative language has not evolved far enough to handle" (p.289, emphasis added). Peters' (1977, 1983) position is more ambiguous. She does mention the possibility of mature speakers using formulaic speech as a shortcutting device (1983, p. 3). However, elsewhere she takes the same stance as Krashen and Scarcella: "I will suggest that items in the lexicon are subject to analysis by the rules as they are induced and those items that yield to such analysis may lose their status as unitary items of storage"(1983, p. 15). From these accounts, it is easy to detect their assumption that formulaic utterances are a device that beginning language learners use to compensate for their lack of creative language and that once their

propositional language is on its way, formulae are no longer needed.

I, on the other hand, propose that formulaic utterances do not have to lose their status as single items in the storage even after their internal structure has been analyzed, as long as they still serve to save processing energy in sentence production. In other words, the most important defining character of such a formula is not that its internal structure is not known to the speaker but that it is stored as a unit in the lexicon and is retrieved as a whole whenever it fits the concept that s/he wants to express. When a speaker produces a sentence, there are a number of things s/he has to attend to other than the use of correct grammar. Given the fact that "humans are limited-capacity information processors" (McLaughlin, Rossman, and McLead 1983, p.137), it is logical to expect that they use whatever device available to save processing time and energy. If one aspect of production does not require much attention, the speaker can afford to pay more attention to other tasks. The use of formulaic utterances is one of such useful devices; regardless of whether their internal structure is perceived or not, if they are stored as single units in the lexicon, they can be retrieved readily, saving energy to construct the same structures from scratch.

Several researchers have pointed out the importance of formulaic expressions in the adult native speakers' speech (Pawley and Syder 1983, Vihman 1982, Gleason 1982, Widdowson 1989). According to Pawley and Syder (1983), in order to select native-like expressions among perfectly grammatical combinations which contain many non-native-like expressions and achieve native-like fluency, speakers must depend quite heavily on the use of formulaic utterances. They note that "native speakers do *not* exercise the creative potential of syntactic rules to anything like their full extent" (original emphasis, p. 193). Their view goes against the often taken-for-granted primacy of the 'principle of parsimony': far from minimizing the amount of description of the lexicon, it in fact promotes an enormous amount of redundancy in it. If these prefabricated sequences are known as units and also can be analyzed syntactically, they must be registered at least twice in the lexicon (Pawley and Syder 1983). However, I agree with Pawley and Syder that what is important is not how economical the description of competence is but how well it fits the psychological reality of linguistic knowledge that human beings possess. If we know

language in a fairly redundant form, the description of the knowledge should reflect that.

Indeed, much of current theory of language production is biased by the principle of parsimony. When Krashen and Scarcella (1978) and Peters (1977, 1983) maintained that formulaic speech eventually yield to more creative language, they were also influenced by the principle. They simply did not see the possibility of both existing at the same time; one or the other had to go. Since the importance of creative language in proficient speakers is indisputable, they had no choice but to discard formulaic speech: an unwarranted bias that is explicitly pointed out by Gleason (1982):

> We have in recent years become so enthralled with the admitted power of generative systems, that memory, as an important process, and the possibly vast store of memorized units we each call upon everyday, have somehow fallen into disrepute . . . . second language learners begin not so much with generative systems as with chunks, prefabricated routines, or unopened packages, as they have been called . . . The importance of routines in language acquisition, in second language learning, and in the everyday use of nonexceptional speakers has yet to be recognized. (p. 355)

If native speakers use formulaic speech quite extensively, there is no reason to posit that intermediate- or advanced-level L2 leaners, whose syntactic and lexical knowledge has developed considerably, do not use it. In fact, It has been suggested that L2 learners use more formulaic utterances than L1 leaners (Hatch 1972). There are a number of reasons for this. First, L2 learners are cognitively more mature than L1 leaners and therefore are capable of memorizing longer sequences (Hatch 1972). In my observation, after five months of exposure to naturalistic as well as classroom English, Philip, 10-year-old Bulgarian boy, who made the most extensive use of formulaic utterances among all the children in my study, came up with such long sequences as Something strange going on here, you know? and Here is you chair, man, come and get it. Second, with the greater semantic development, they have a greater need to communicate their thoughts (Hakuta 1974). And third, L2 learners in a naturalistic setting are compelled to fare with native speakers and are under constant pressure to manipulate the language as fast as they can. If they do not speak fast enough, native speakers will soon interrupt. Thus, L2 learners have more

incentives as well as abilities to use formulaic speech than native speakers (Krashen and Scarcella 1978). What is important to note is that these incentives are just as strong for intermediate and advanced L2 learners as for beginners. As learners' L2 develops, their own expectations of what they should be able to do in L2 change together with the expectations of people around them. Very advanced learners, for instance, are expected to produce L2 sentences just as fluently and coherently as native speakers, despite the fact that it is still an L2 for them. They are thus just as much in need of energy-saving devices for sentence production as beginning learners, which enables them to allocate limited processing capacity to other tasks.

Empirical support

Logically, then, there is nothing strange about L2 learners using formulaic utterances whose internal structure they know. Is there any empirical evidence in SLA data to show that this is the case? Although there has been no empirical study which has dealt directly with this issue, there is a little information here and there that, when put together, points to that direction.

First of all, my own data suggest that a learner's stock of memorized formulae and syntactic and lexical development are not as independent from each other as have been considered. In fact, there seems to be a closer correlation between the syntactic, lexical sophistication of a child's formulaic utterances and and his/her productive abilities. Sally, an eight-year-old Sri Lankan girl, was far behind everyone else in her syntactic development. In the last session of my observation (i.e. after seven months), she was still not able to insert verbs in her propositional speech in most occasions. Here are some examples of her creative construction:

(5) Me good girl, Ron bad boy.

(6) Me tree. [meaning "I have three".]

(7) Me finished. [meaning " I am finished". Finished is one of her formulae.]

(8) Grandfather is old. [this was the only occasion that she managed to insert the copula in her creative speech in this session.]

Her repertoire of formulae, on the other hand, was also very limited: What's this?, What's that?, Oh my god, You shut up, Finished, and Hurry up. These relatively simple expressions were accessible to her because their meaning was easy enough to guess from the context without much syntactic and lexical knowledge.

In contrast, Ron, eight-year old Rumanian boy, who was the most successful English learner in my study, had considerably more complex formulae at hand:

(9) I'm gonna kill you next time.
(10) Look what you did.
(11) You are in big trouble.
(12) You know song that everybody sings?

The last example is particularly interesting. This utterance was clearly a formula since he pronounced it without any pause and repeated exactly the same phrase several times that day. It sounded like a recitation. Every time he used it, it had nothing to do with the activity that he was engaged in. What is interesting about this utterance is its relation to another utterance he had made two weeks before, which was clearly a *rule-based* production:

(13) Hunter is the one man who shoots fox.

It was the first time that I observed his use of a relative clause. He said this in reference to the picture book his ESL teacher was reading, in which a hunter appeared. What is significant here is that he acquired a formulaic expression containing a relative clause *after* he started to use relative clauses productively. In more general terms, some formulaic utterances are so complex that the learner must be able to parse them at least partially in order to understand their meaning. That is, in order to acquire and use fairly complex expressions as formulae, (partial) knowledge of their internal structure is a prerequisite. As the learner's syntactic and lexical development proceeds, s/he becomes capable of understanding more complex expressions and store some of them as units. There is always a gap between one's receptive and productive abilities; expressions that learners hear and understand but cannot produce themselves may be the prime candidate for formulaic utterances. This is consistent with my position that formulae may remain

as unitary items in the lexicon, regardless of the learner's accessibility to their syntactic structure.

Second, other researchers' data also suggest the existence of formulaic expressions as units after the syntactic analysis. One of the strategies that Clark's son, Adam, used was to include the whole of his previous utterances as a part of the next utterance:

(14) Baby Ivan have a bath, let's go see baby Ivan have a bath.

(15) Adam: Mummy you go.
Mother: Where?
Adam: Mummy you go swings. (Clark 1974, p. 2)

The pattern that Adam generated creatively in the first utterance was stored as a unit temporarily and used as a chunk in the second utterance. This way he did not have to construct that part of the second sentence from scratch and thus could attend to other parts. Since in this case, the formula in the second sentence was originally his own rule-governed production, its syntactic structure was obviously known to him. Note that formulaic utterances can be quite temporary constructs, as in this case, something to be forgotten as soon as they serve the intended purpose.

Another example that Clark gives lends further support to my argument. At one point, Adam was able to modify his idiosyncratic structure according to his mother's positive input:

(16) Adam: My shoe on a polish. [Trying to put polish on his shoes.]
Mother: The polish is on your shoes.
Adam: Polish on my shoes. (1974, p. 8)

However, when he wanted to insert another element into the sentence, he reverted to his old, more familiar structure:

(17) My shoes on a *brown* polish. (1974, p. 9)

A very similar phenomenon is observed in Hakuta (1974). His subject, Uguisu acquired these are as one of her prefabricated patterns in the second month. In the

fifth month, when her number agreement was getting established, she was observed to revert to her old formula in one occasion:

(18) These is for . . . these are for big person like my, I. (p. 292)

What these two examples suggest is that when, for one reason or another, processing cost exceeds the learner's capacity, s/he tends to resort to prefabricated expressions to take off some of the burden, as McLaughlin et al. (1983) notes: "The execution of new skills is costly in terms of workload involved and will occur only when other tasks and cognitive demands are minimized" (p. 145). In order to revert to formulaic utterances, however, the learner must keep them in the lexicon even after their internal structures have become clear.

If we accept the view that the internal structures of some formulaic expressions are known to the learner, we may be able to see far more of them at work in a L2 learner's production than we have so far been aware of. Take Uguisu's production of question forms in Hakuta (1974) for example. Among the list of numerous question forms that Uguisu used, Hakuta recognizes only one prefabricated pattern, i.e. do you. This is because these two words appeared together all the time from the very beginning; no other combination was used. Hakuta is probably correct in inferring that initially do you was not segmented and was learned as a question marker. However, a glance at his Table 3, part of which is reproduced in Table 1 below, clearly shows that there are also several other questions or parts of a question that are used repeatedly in the same form during a particular period of acquisition:

Table 1: examples of the question forms that Uguisu produced (taken from Hakuta 1974, p. 294 Table 3)

Month 4	Where did you get that?
Month 5	What did she say?
	What did you say?
	What did you say?
Month 6	What did you do?
	What did you say?
Month 8	Do you saw these peppermint?
	Do you saw some star eye?

	Do you saw some star eye?
	Did everybody saw some blue hairs?
Month 10	Why did you do that?
	Why did you get this?
	Why did you go to a hospital?
	Why did you draw?
Month 11	Why did you put this?

Quite visibly there is what may be called the 'expression of the month' in Uguisu's development of question forms. It is equally apparent that she is capable of segmenting the elements of these expressions since she sometimes combines them with other words. For instance, what did you is used repeatedly in month 5 and 6, but already in month 4, she combined did you with where, which indicates that what and did you were syntactically separable for her by the time she used what did you. Similarly, why did you became a formula for her in month 10 and 11, although by then she was capable of segmenting all the three words in the expression. Another example, do you saw, is particularly interesting in this respect. Since this could not have happened in the input, it must have been her own creative speech turned into a chunk just like Adam's case (see examples 14 and 15 above).

Another phenomenon which may be better explained in term of formulaic utterances as an "energy saver" in speech production is overgeneralization in the learner's interlanguage. Again, Hakuta's data (1974) is pertinent. In the first stage, up to month 8, Uguisu inverted the subject and the verb in the embedded wh-questions most of the time:

(19) I don't know where is it.
(20) I don't know where is your house.
(21) You know where is my house. (p. 296 take from Table 4)

This has traditionally been considered as a case of overgeneralization: sub-aux inversion in the main clause is incorrectly applied to the subordinate clause. In month 8, Uguisu goes onto the second stage of development, using the copula before and after the subject:

(22) You will see where is your house is.

(23) I don't know where is the telephone number is.
(24) I don't know where is the woods is. (Ibid.)

Then in the last stage in month 9, the copula finally remains uninverted:

(25) I know where it is.
(26) I don't know where the bathroom is. (Ibid.)

Overgeneralization can account for the first stage: the learner assumes that what is applicable in the main clause is also true for the subordinate clause. However, it fails to explain the second stage of development such as (22) (23) and (24) since there is another rule operating that contradicts the previous hypothesis. If the learner assumed that the copula must be inverted, then why would she put it after the subject again? It seems to me that this development of embedded wh-questions can be explained more reasonably if we regard where is as a formula.

In the first stage, Uguisu invariably inverts the subject and the copula because where is is a fixed formula for her. It is reasonable to assume so since simple, unembedded wh-questions, in which sub-aux inversion is obligatory, are normally mastered before they are embedded in another clause. While learning to produce simple question forms, where is must have become a formula for her. Then there is nothing to prevent her from using this formula in the subordinate clause as long as she does not know the grammatical rule that prohibits it. Although it is grammatically incorrect, the meaning of the sentence is perfectly understandable.

Up to this point, there is very little to distinguish this "formula appropriation" from overgeneralization. However, the former explanation can hold in the second stage of development while the latter cannot. According to the "least effort principle" (Pawley and Syder 1983), Uguisu retrieves where is as a formula in order to save the processing workload; indeed, at this stage where is must be tied so strongly together in her memory that it is difficult for her to separate the two words without spending some energy for it. On the other hand, she is also aware of a new grammatical rule, although not as explicitly as to make her refrain from using the formula. So according to this rule, she inserts the copula after the subject. At this stage, both competing strategies--formula appropriation and rule observation--are at work but

neither is strong enough to negate the other. After a while, however, the rule becomes sufficiently strong to keep Uguisu from using the formula in this situation, thus yielding the correct form as in (25) and (26).

Brown and Hanson's (1970) case, which I quoted above, might be very similar. When prefabricated expressions What's that? and What are you doing? persisted in embedded clauses while other wh-questions are uninverted, they explained that these expressions become rigid enough in the learner's lexicon so that they become immune to structural reanalysis for a while. This is quite possible. However, another equally plausible explanation is that, even after they have been subject to reanalysis, they might remain in their prefabricated forms as a convenient energy saver until the learner becomes sufficiently aware of the grammatical rules that prevent the use of such expressions. In other words, Brown and Hanson's learners might have been capable of analyzing these expressions; they nonetheless used them as a short-cut to reduce processing energy until they became aware that they could not do so in this context.

In sum, formulaic utterances that are no longer useful or detrimental to rule development are either discarded entirely (although I think this is unlikely) or get certain constraints attached to them as to in which linguistic contexts they can be used. Thus where is for Uguisu became segmented in the embedded clauses but it might remain as a useful formula in the main clause questions. On the other hand, there are other numerous formulaic expressions which are highly useful and which do not disturb rule formation. These are the formulae that even native speakers use. Pawley and Syder (1983) maintain that adult native speakers use tens of thousands of them. Then formulaic speech is not a strategy that is used only at the beginning of the language development but continues to be used throughout the process and even at the final stage and ever after. How are you?, I am sorry, I should have known etc., etc. are just a few examples of such stable formulae.

Summary and theoretical implications

Thus far I have argued that the most important defining character of formulaic utterances is *not* that their internally structure is unknown to the learner but that they are stored as unitary items in the lexicon. A summary of what I have said so far

may be helpful for the reader to clarify the most important points:

(1) Even after the internal structures of formulaic expressions become clear to the learner, these expressions can remain as 'chunks' in the lexicon as long as they serve to save processing cost in speech production.

(2) Learners thus might resort to formulaic utterances even when they are vaguely aware that it is grammatically wrong to do so. This is likely to happen when they have to introduce new items which require much processing energy and the overall workload becomes more than they can handle.

(3) When learners have acquired grammatical rules that prohibit the use of a formula in certain contexts, they will avoid it in these contexts. However, as long as it is still useful in other contexts, it may remain as a formula. In other words, formulaic expressions that are not "disapproved" by other rules remain.

(4) Formulaic utterances play a crucial role not only in the beginning stage of language development but also in the later stages and even in the native speaker's speech production.

(5) Formulae range from fairly permanent to short-lived ones.

What theoretical implications do formulaic utterances defined as above have in language development? I will mention two that I am currently aware of. First of all, interlanguage development has so far been characterized in terms of grammatical rule formation and vocabulary increase. However, if one takes into account the role of formulaic utterances in language development and also the fact that native speakers possess a large set of such formulae, IL development can also be defined in terms of the acquisition and selection of formulaic utterances. Learners start out with idiosyncratic formulae, discarding those that are mutually exclusive with newly acquired rules and storing syntactic and lexically more complex expressions as their knowledge in syntax and vocabulary increases, until their set of formulaic utterances converges on that of native speakers. Admittedly, it is difficult to define exactly what a native speaker's formulaic speech inventory consists of. A person's

formulaic utterances range from what Peters (1983) calls "cultural formulas", i.e. expressions that are shared as units in a particular speech community, to "idiosyncratic formulas", i.e. patterns that only s/he uses as fixed forms. Thus no two speakers have exactly the same set of formulaic expressions. However, in so far as there is such thing as "cultural formulas", learners must acquire them in order to be judged competent speakers.

This last point leads me to another theoretically important implication. One of the major problems of language acquisition has been the lack of negative evidence (White 1985, 1987, Pinker 1984): how can a learner tell what s/he is saying is wrong if there is nothing in the input to suggest that? However, if we consider the degree to which we rely on formulaic expressions, the learner might get the negative evidence from the lack of the item in the input. If the concept that learners express in their idiosyncratic fashion is represented in another form in the input, then they might infer that their own form is *conventionally* wrong and switch to the form that exists in the input. Expressions that we have never heard before do tend to sound strange. Language use may not be as productive as Chomsky has thought, although creative aspects of language development and language use are indisputably significant. There is a great deal of habit formation in our use of language.

References

Brown, R. and Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In Hayes, J. R. (Ed.), Cognition and the development of language (pp. 11-54). New York: John Willey & Sons , Inc.

Clark, R. (1974.) Performing without competence. Journal of Child Language, 1, 1-10.

Gleason, J. B. (1982). Converging evidence for linguistic theory from the study of aphasia and child language. In Obler, L. K. and Menn, L. (Eds.), Exceptional language and linguistics (pp.347-356). New York: Academic Press.

Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition. Language Learning, 24 (2), 287-297.

Hatch, E. (1972). Some studies in language learning. UCLA Workpapers in Teaching English as a Second Language Acquisition, 6, 29-36.

Krashen, S. and Scarcella, R. (1978). On routines and patterns in language acquisition and performance. Language Learning, 28 (2), 283-300.

McLaughlin, B., Rossman, T. and McLead, B. (1983). Second language learning: an information-processing perspective. Language Learning, 33 (2), 135-157.

Pawley, A. and Syder, H. (1983). Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In Richard, J. C. and Schmidt, R. W. (Eds.), Language and Communication (pp. 191-226). London: Longman.

Peters, A. M. (1977). Language learning strategies: does the whole equal the sum of the parts? Language, 53, 560-573.

Peters, A. M. (1983). The units of language acquisition. Cambridge: Cambridge University Press.

Pinker, S. (1984). Language learnability and language development. Cambridge, MA: Harvard University Press.

Vihman, M. M. (1982). Formulas in first and second language acquisition. In Obler, L. K. and Menn, L. (Eds.), Exceptional language and linguistics (pp.261-284). New York: Academic Press.

White, L. (1985). The acquisition of parameterized grammars: subjacency in second language acquisition. Second Language Research, 1 (1), 1-17.

White, L. (1987). Against comprehensible input: the input hypothesis and the development of second-language competence. Applied Linguistics, 8 (2), 95-110.

Widdowson, H. G. (1989). Knowledge of language and ability for use. Applied Linguistics, 10 (2), 129-137.

Wong Fillmore, L. (1976). The second time around: cognitive and social strategies in second language acquisition. Unpublished Doctoral dissertation, Stanford University.

Wong Fillmore, L. (1979). Individual differences in second language acquisition. In Fillmore, C. J., Kempler, D. and Wang, W. S-Y. (Eds.) Individual differences in language ability and language behavior (pp. 203-228). New York: Academic Press.